DOCUMENT RESUME

ED 120 206 95 TM 005 165

AUTHOR Godbout, Robert C.; And Others
TITLE Analysis of Unbalanced Data.
INSTITUTION Texas Univ., Austin. Research and Development Center for Teacher Education.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE [Apr 75]
CONTRACT NIE-C-74-0088
NOTE 34p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30-April 3, 1975)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage
DESCRIPTORS Analysis of Variance; *Data Analysis; *Educational Research; Hypothesis Testing; Multiple Regression Analysis; Research Design; *Research Methodology; *Research Problems; *Statistical Analysis
IDENTIFIERS *Unbalanced Data

ABSTRACT

The present paper contributes to research methodology as practiced in the field by providing practitioners with a concise statement of the problems in analyzing unbalanced designs, and by clarifying the conditions under which the analysis alternatives are most appropriate. The purposes of the present paper are: (1) to examine the analysis alternatives which have been suggested for unbalanced data, (2) to determine the interpretation appropriate to each type of analysis, and (3) to suggest which of the techniques are appropriate for different types of research questions. The above concepts are illustrated with data from typical research studies. (Author)

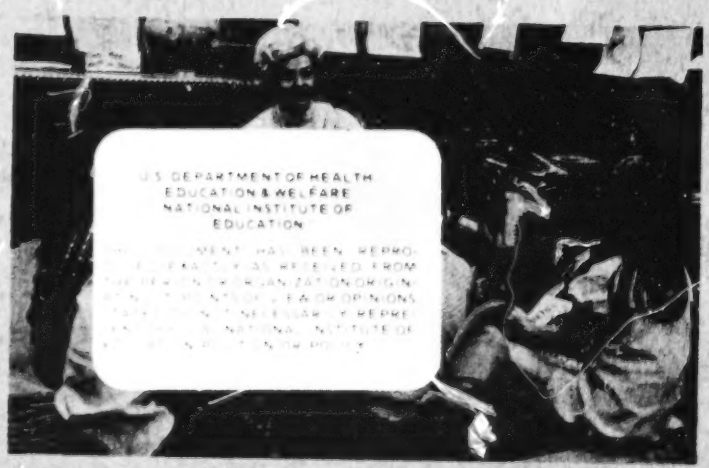
****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED120206

a UTR&D report

EVALUATION OF TEACHING PROJECT

TM005 165



2

THE RESEARCH AND DEVELOPMENT CENTER FOR TEACHER EDUCATION • THE UNIVERSITY OF TEXAS AT AUSTIN

The *Evaluation of Teaching Project*—one of four projects at the Research and Development Center for Teacher Education—has as its mission to develop materials and strategies for teacher training, research and evaluation. The goals of the *Evaluation of Teaching Project* are to develop (1) a conceptual framework for the evaluation of teaching, (2) a sourcebook of validated teacher evaluation and research instruments, and (3) strategies for the evaluation of teacher trainee programs. These goals are being carried out with funds from the National Institute of Education.

In the process of meeting its objectives, the *Evaluation of Teaching Project* conducts systematic research in teacher behavior for the purpose of validating instruments and identifying characteristics of effective teaching. The following report describes one facet of this research. A complete listing of studies in this report series is available by writing to the Evaluation of Teaching Project, R&D Center for Teacher Education, University of Texas at Austin, 78712.

Analysis of Unbalanced Data

Robert C. Godbout, Gary D. Borich

and Lee H. Poynor

The University of Texas at Austin

Paper to be presented at the annual meeting of the American Educational Research Association, Washington, D. C., April, 1975.

The research reported herein was supported in part by National Institute of Education Contract NIE-C-74-0088, Research and Development Center for Teacher Education. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by those offices should be inferred.

Analysis of Unbalanced Data

Robert C. Godbout, Gary D. Borich and Lee H. Poynor

The University of Texas at Austin

Researchers in experimental disciplines can randomly divide subjects into groups which then receive different treatments. Provision can easily be made to assign equal numbers of subjects to the different groups. Given such an equal-n or balanced design, inferential statistical techniques which are relatively simple to calculate and yield clear interpretations are available. In contrast, the educational researcher is rarely afforded the luxury of easy experimental manipulations. Such manipulations may be either impossible or ethically questionable. The educational researcher, wishing to study the effect of characteristic X, must usually draw a sample of subjects who already possess X and compare this with a sample who do not possess X. If the absence and presence of X are not of equal frequency in the population being sampled, then unequal numbers of subjects possessing and not possessing X will likely be obtained. In general, field samples lead to unequal-n or unbalanced designs. Given an unbalanced design, the normal statistical techniques do not provide clear interpretations. This is because the different effects in such an unbalanced design are confounded. Special techniques have been derived to eliminate such confounding, but it is unclear which of these techniques should be used for a particular research question. It is also unclear exactly what interpretations are appropriate to a given technique. The purposes of the present paper are to examine the different types of analyses suggested for unbalanced data, to determine which interpretation is appropriate to each, and to suggest which techniques are appropriate for different types of research questions.

No attempt has been made in the present paper to provide the detailed mathematical derivations and calculation schemes for the various solutions. Such matters have been more than adequately covered in numerous other sources (e.g., Kirk, 1968; Overall and Spiegel, 1969; Searle, 1971; Winer, 1971; Rao, 1973; Marks, 1974). The present intent is to examine the different unbalanced data techniques from the point of view of the educational researcher in the field. A researcher with the practical need of finding an appropriate technique for analyzing an unbalanced design has little concern for subtleties in mathematical derivations. Nor has he much interest in the details of calculating the desired statistics. The advent of large computers and statistical program packages has freed the researcher from most calculation drudgery. This is not to say that a researcher can mindlessly select a computer program capable of analyzing unbalanced data, feed his data into a computer with that program, and then report the answers provided on the printout. The answers are meaningless unless one knows what questions they relate to. Different statistical techniques ask different questions of the data--i.e., test different hypotheses and require different interpretations.

There are four basic steps in the analysis of a data set: (1) determination of the relevant questions to be asked; (2) selection of an analysis technique appropriate to those questions; (3) analysis of the data; and (4) interpretation of the results. While the first two steps are often omitted, all four steps are necessary for meaningful research. It is essential that the researcher know the questions posed by different analysis techniques and the interpretations appropriate to these techniques. Our intention is to attempt to provide such information for the various techniques for analyzing unbalanced data.

For the present purposes, it is unnecessary to consider complex designs. The interpretations appropriate to the different analysis techniques can be illustrated with quite simple designs. The conclusions arrived at on the basis of such simple designs are easily generalized to more complex designs. As a result, the designs considered in the present paper are usually two-factor designs with all effects considered to be fixed effects. The remainder of the paper is devoted to a presentation of (1) techniques based upon the traditional analysis of variance approach, (2) Jennings' (1967) regression analysis approach, and (3) other approaches to unbalanced data analysis. The final section involves a comparison of the different methods.

Analysis of Variance Methods

Analysis of variance techniques were originally derived for the analysis of balanced data. However, many writers (Winer, 1971; Searle, 1971; Rao, 1973; Marks, 1974) have shown that the basic logic of analysis of variance can be extended to the domain of unbalanced data. As the term "analysis of variance" suggests, the method is concerned with partitioning variance on the dependent variable (criterion) into components attributable to the different independent variables (predictor variables) in the design.[1] The amount of variance attributable to an independent variable is taken as a measure of the effect of that variable. That portion of variance not accounted for by the independent variables is the error variance. The statistical significance of a variance component is evaluated through comparison with the error variance.

The most widely used method for conceptualizing analysis of variance is in terms of the general linear model. The general linear model for a two-variable design is:

---

[1]Technically, the criterion sum of squares is partitioned rather than the variance.

$$X_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}. \qquad (1)$$

The symbols in (1) are: X_{ijk}, the dependent variable (criterion) score for the k^{th} individual in the ij^{th} cell; μ, the grand mean; α_i, the effect of the i^{th} row (i^{th} level of the A variable); β_j, the effect of the j^{th} column (j^{th} level of the B variable); α_{ij}, the interaction effect for the ij^{th} cell; and ε_{ijk}, the error for the k^{th} individual in the ij^{th} cell. The error component refers to the unpredictable part of a score. Error is assumed to be random and it follow that $\sum_k \varepsilon_{ijk} = 0$ for all i and j.

As can be seen from expression (1), each score is viewed as the sum of several components. The inclusion of the grand mean (μ) allows us to view the other components as deviations from that grand mean. Thus, the value of α_1 may be -10 and the scores for all subjects who experience the first level of the A variable are visualized as being decreased by 10 because of that experience. The value of α_2 may be +8 and the scores for all subjects who experience the second level of the A variable are conceptualized as being increased by 8 because of that experience. The set of elements α_i for all i reflect the <u>main effect</u> of variable A. A main effect is an effect which is constant for <u>all</u> subjects under consideration. It is the manner in which all individuals are uniformly affected by a condition or treatment.

In expression (1), the set of $\alpha\beta_{ij}$ terms for all i and j reflect the <u>interaction</u> of variables A and B. The interaction refers to effects of the independent variables which are not constant for all individuals--i.e., effects not ascribable to the main effects. The interaction of variables A and B involves the residual predictable variation not ascribable to the main effects of A and B.

Dummy Variable Coding

Solutions to analysis of variance problems can be obtained by first estimating the parameters in the general linear model and then determining sums of squares and variance components on the basis of these estimates (e.g., Rao, 1973). However, simpler methods are available (e.g., Overall and Spiegel, 1969; Kerlinger and Pedhazur, 1973). These simpler methods entail defining a set of dummy variables corresponding to the analysis of variance parameters. If the set of dummy variables is defined in an appropriate manner, then the required analysis of variance can be performed by examining correlations between these dummy variables and the criterion variable.

Let Z(p) be a dummy variable corresponding to the analysis of variance parameter p. For a 2x2 design, the general analysis of variance model has nine parameters--μ, α_1, α_2, β_1, β_2, $\alpha\beta_{11}$, $\alpha\beta_{12}$, $\alpha\beta_{21}$, and $\alpha\beta_{22}$--and a dummy variable set for such a design will have nine members--$Z(\mu)$, $Z(\alpha_1)$, $Z(\alpha_2)$, etc. We now turn our attention to two different methods of coding dummy variables: effect coding and contrast coding.[2]

Effect coding. Effect coding for a 2x2 design is illustrated in Table 1.

---

[2]The reader should be cautioned that the labels for dummy variable coding systems are not consistent in the literature. For example, the coding system called effect coding in the present paper has been called dummy variable coding (Searle, 1971) and indicator variable coding (Winer, 1971). The system called contrast coding in the present paper has been referred to as dummy variable coding (Overall and Spiegel, 1969), contrast coding (Cohen, 1968), and orthogonal coding or effect coding (Kerlinger and Pedhazur, 1973). Care should be taken to correctly identify the systems implied by the present labels.

Table 1

Effect Coding for a 2x2 Design

Group	$Z(\mu)$	$Z(\alpha_1)$	$Z(\alpha_2)$	$Z(\beta_1)$	$Z(\beta_2)$	$Z(\alpha\beta_{11})$	$Z(\alpha\beta_{12})$	$Z(\alpha\beta_{21})$	$Z(\alpha\beta_{22})$
A_1B_1	1	1	0	1	0	1	0	0	0
A_1B_2	1	1	0	0	1	0	1	0	0
A_2B_1	1	0	1	1	0	0	0	1	0
A_2B_2	1	0	1	0	1	0	0	0	1

The columns of Table 1 represent the various dummy variables and the rows represent each of the four groups in the design. Subjects in the A_1B_1 cell are assigned the Z(p) values in the first row; subjects in the A_1B_2 cell are assigned the values in the second row; etc. Note that subjects experiencing condition A_1 receive a value of 1 for the dummy variable corresponding to α_1 and a 0 for the dummy variable corresponding to α_2. For subjects experiencing A_2, the value of $Z(\alpha_2) = 1$ while the value of $Z(\alpha_1) = 0$. In general, $Z(p) = 1$ for all subjects experiencing the condition corresponding to p, while $Z(p) = 0$ for all other subjects. Effect coding has been suggested by Searle (1971) and Winer (1971).

Contrast coding. This method of coding dummy variables has been suggested by Cohen (1968), Overall and his coworkers (Overall and Spiegel, 1969; Overall and Klett, 1972), and Kerlinger and Pedhazur (1973). Table 2 illustrates contrast coding for a 2x2 design.

Table 2

Contrast Coding for a 2x2 Design

Group	$Z(\mu)$	$Z(\alpha_1)$	$Z(\alpha_2)$	$Z(\beta_1)$	$Z(\beta_2)$	$Z(\alpha\beta_{11})$	$Z(\alpha\beta_{12})$	$Z(\alpha\beta_{21})$	$Z(\alpha\beta_{22})$
A_1B_1	1	1	0	1	0	1	0	0	0
A_1B_2	1	1	0	-1	0	-1	0	0	0
A_2B_1	1	-1	0	1	0	-1	0	0	0
A_2B_2	1	-1	0	-1	0	1	0	0	0

Contrast coding is characterized by the fact that there are only as many dummy variables as there are groups. Note that $Z(\mu)$, $Z(\alpha_1)$, $Z(\beta_1)$, and $Z(\alpha\beta_{11})$ are the only variables in Table 2 which possess nonzero values. The five other variables are always 0 and need not be considered. In contrast coding, each variable carries a contrast or comparison among the groups. For example, $Z(\alpha_1)$ involves comparison of groups A_1B_1 and A_1B_2 vs. groups A_2B_1 and A_2B_2. $Z(\alpha_1)$ contrasts subjects experiencing A_1 with subjects experiencing A_2--i.e., $Z(\alpha_1)$ carries the information relevant to the A main effect. The variable $Z(\beta_1)$ involves comparison of groups experiencing B_1 with groups experiencing B_2 and carries the information about the B main effect. The interaction variable, $Z(\alpha\beta_{11})$, is obtained as the product of the two main effect variables. In general, $Z(\alpha\beta_{ij}) = Z(\alpha_i) \cdot Z(\beta_j)$.

To illustrate further properties of contrast coding, consider Table 3.

Table 3

Contrast Coding for a One-Way Design with Six Levels

Group	$Z(\mu)$	$Z(\alpha_1)$	$Z(\alpha_2)$	$Z(\alpha_3)$	$Z(\alpha_4)$	$Z(\alpha_5)$
A_1	1	1	0	0	0	0
A_2	1	-1	1	0	0	0
A_3	1	-1	-1	1	0	0
A_4	1	-1	0	-1	1	0
A_5	1	-1	0	0	-1	1
A_6	1	-1	0	0	0	-1

In Table 3, there are six dummy variables corresponding to the six groups. $Z(\alpha_1)$ carries a contrast involving group 1 vs. the other five groups. $Z(\alpha_2)$, $Z(\alpha_3)$, $Z(\alpha_4)$, and $Z(\alpha_5)$ involve contrasts of 2 vs. 3, 3 vs. 4, 4 vs. 5, and 5 vs. 6, respectively.

The coding in Table 3 will lead to a valid analysis of variance. An important feature of the set of contrasts in Table 3 is that they are not orthogonal. Workers familiar with techniques involving orthogonal comparisons in analysis of variance (see Edwards, 1968, or Winer, 1971) may have assumed that the present contrasts need be orthogonal. This is not the case. While the contrasts may be orthogonal (as they were in Table 2), they need not be.

As Cohen (1968) has pointed out, a set of contrast-coded dummy variables will yield a valid analysis of variance if the following two conditions are met: (1) there are as many contrasts as there are groups and (2) no dummy variable gives a multiple correlation coefficient of unity with the other dummy variables. The method of contrast coding is quite flexible and there is no standard method of arriving at the codes. Of course, if the set

of dummy variables does not involve contrasts relevant to main effects and interaction, then they cannot be employed to ask traditional analysis of variance questions. In addition, some methods of writing contrast-coded dummy variables yield intermediate results of practical value. For a particularly practical method of contrast coding see Overall and Spiegel (1969) or Overall and Klett (1972).

Partitioning Sums of Squares

Performing an analysis of variance by analyzing dummy variables will now be considered. Given an appropriate set of dummy variables, one can calculate multiple correlation coefficients for the criterion variable with various subsets of these dummy variables. The squares of such multiple correlation coefficients can be used to determine the proportions of the criterion variance attributable to the different analysis of variance sources. Consider a two-factor design. Let $R(\alpha, \beta, \alpha\beta)$ be the multiple correlation for the entire set of dummy variables with the criterion; $R(\alpha, \beta)$, the multiple correlation for the subset of dummy variables involving the two main effects; $R(\alpha)$, the multiple correlation for the set of A main effect dummy variables; $R(\beta)$, the multiple correlation for the set of B main effect dummy variables; etc.

The sum of squares attributable to the A main effect can be calculated in four different ways:

$$\begin{aligned} SS_A &= SS_T[R^2(\alpha)] \\ SS_{A/B} &= SS_T[R^2(\alpha, \beta) - R^2(\beta)] \\ SS_{A/AB} &= SS_T[R^2(\alpha, \alpha\beta) - R^2(\alpha\beta)] \\ SS_{A/B,AB} &= SS_T[R^2(\alpha, \beta, \alpha\beta) - R^2(\beta, \alpha\beta)]. \end{aligned} \tag{2}$$

The term SS_T represents the total, mean corrected, sum of squares for the criterion variable. SS_A represents the total (unadjusted) sum of squares for the A main effect; $SS_{A/B}$, the sum of squares for the A main effect with the B effect covaried or partialed out; $SS_{A/AB}$, the sum of squares for the A main effect with the interaction covaried or partialed out; and $SS_{A/B,AB}$, the sum of squares for the A main effect with the B effect and the interaction covaried or partialed out.

Expressions similar to (2) can be written for the sum of squares for the B main effect:

$$\begin{aligned} SS_B &= SS_T[R^2(\beta)] \\ SS_{B/A} &= SS_T[R^2(\alpha, \beta) - R^2(\alpha)] \\ SS_{B/AB} &= SS_T[R^2(\beta, \alpha\beta) - R^2(\alpha\beta)] \\ SS_{B/A,AB} &= SS_T[R^2(\alpha, \beta, \alpha\beta) - R^2(\alpha, \alpha\beta)]. \end{aligned} \tag{3}$$

Finally, a single expression can be written for the interaction sum of squares:

$$SS_{AB/A,B} = SS_T[R^2(\alpha, \beta, \alpha\beta) - R^2(\alpha, \beta)]. \tag{4}$$

Expressions for $SS_{AB/A}$, $SS_{AB/B}$, and SS_{AB} have not been written. The residual character of interaction has been previously stressed. This residual character dictates that the interaction sum of squares be of the form, $SS_{AB/A,B}$.

Assuming a fixed effects model, the variance component corresponding to any of the sums of squares in (2), (3) or (4) can be tested for significance with the variance component corresponding to:

$$SS_{error} = S_T[1 - R^2(\alpha, \beta, \alpha\beta)]. \tag{5}$$

Let F(A) be the F-ratio used to test the significance of the variance component corresponding to SS_A; F(A/B), the F-ratio corresponding to $SS_{A/B}$; F(A/AB), the F-ratio corresponding to $SS_{A/AB}$; and F(A/B, AB), the F-ratio corresponding to $SS_{A/B,AB}$.

At this point, it is necessary to reconsider the two types of dummy variable coding--effect coding and contrast coding. Given an unbalanced design and given that we have used contrast coding, different values will likely be obtained for SS_A, $SS_{A/B}$, $SS_{A/AB}$, and $SS_{A/B,AB}$. F(A), F(A/B), F(A/AB), and F(A/B, AB) will also likely differ, and these four F-ratios will reflect tests of four different hypotheses. A different state of affairs arises if effect coding is employed. When effect coding is employed, $SS_{A/B,AB} = SS_{A/AB} = 0$ for any design (Searle, 1971, p. 301). In the case of effect coding, F(A/B, AB) and F(A/AB) do not provide tests of any hypothesis. The levels of interaction--i.e., the $Z(\alpha\beta_{ij})$ dummy variables when we use effect coding--cannot serve meaningfully as covariables since such covariables will account for all of the predictable variance, leaving no variance to be accounted for by the main effects. Contrast coding allows the interaction to be covaried for tests of main effects, whereas effect coding does not. Contrast coding, therefore, seems to be a more useful and powerful dummy variable coding scheme than effect coding. The remainder of this discussion of analysis of variance techniques will assume the use of contrast coding.

Classification of Designs

Four different types of designs can be distinguished with regard to hypotheses tested by analysis of variance. These four types of designs are (1) equal-n designs, (2) proportional-n designs, (3) unbalanced (disproportional-n) designs with no empty cells, and (4) unbalanced designs with empty cells. These types of designs will be discussed in turn.

Equal-n designs. Given a design with equal cell frequencies, all main

effects and interactions are mutually independent. This implies for the two-factor case that $SS_A = SS_{A/B} = SS_{A/B,AB} = SS_{A/AB}$. Furthermore, F(A), F(A/B), F(A/B,AB), and F(A/AB) all test the same hypothesis, that hypothesis being:

$$H: \text{ equality of } \alpha_i + \overline{\alpha\beta}_{i.} \text{ for all } i. \tag{6}$$

The symbol $\overline{\alpha\beta}_{i.}$ refers to the mean interaction effect for the i^{th} row (i^{th} level of the A variable). That is,

$$\overline{\alpha\beta}_{i.} = \frac{\sum_j \alpha\beta_{ij}}{j}.$$

In the equal-n case, the residual nature of interaction dictates that

$$\sum_i \alpha\beta_{ij} = 0 \text{ for all } j$$

$$\text{and } \sum_j \alpha\beta_{ij} = 0 \text{ for all } i.$$

If either of these expressions are not true, then variation attributable to a main effect has been incorrectly assigned to interaction--i.e., the interaction has not been treated as a residual source. Thus $\overline{\alpha\beta}_{i.}$ equals 0 and the hypothesis in (6) reduces to equality of the α_i for all i. The analogous hypothesis tested with regard to the B main effect is equality of the β_j for all j. In the case of equal-n data, a clear test of main effects is provided by analysis of variance.

Proportional-n designs. A proportional-n design is one in which the cell frequencies satisfy the following conditions of proportionality:

$$n_{ij} = \frac{n_{i.} n_{.j}}{n_{..}} \text{ for all } i \text{ and } j. \tag{7}$$

The symbols in this expression are: n_{ij}--the number of subjects in the cell created by the intersection of the i^{th} row and j^{th} column, $n_{i.}$--the total number of subjects in the i^{th} row, $n_{.j}$--the total number of subjects in the j^{th} column, and $n_{..}$--the total number of subjects in the design. Expression

(7) is the definition of proportionality appropriate to a two-factor design. Analogous expressions can be written for designs involving more than two independent variables.

In the proportional-n case, main effects are independent, but the unequal frequencies result in correlations between interactions and main effects. Considering a two-factor design, the following relationships hold:

$$SS_A = SS_{A/B} = c_1$$

$$SS_{A/B,AB} = SS_{A/AB} = c_2$$

$$\text{but } c_1 \neq c_2 .$$

Thus, in the proportional-n design, we can arrive at two different sums of squares for the A variable. It follows that two different significance tests can be performed with regard to that variable.

The F-ratios, F(A) and F(A/B), both test the same hypothesis, that hypothesis being:

$$\text{H: equality of } \alpha_i + \frac{1}{n_{i.}} \sum^{j} n_{ij} \alpha\beta_{ij} \text{ for all i.} \qquad (8)$$

The F-ratios, F(A/B,AB) and F(A/AB), test an alternative hypothesis, that hypothesis being:

$$\text{H: equality of } \alpha_i + \overline{\alpha\beta}_{i.} \text{ for all i.} \qquad (9)$$

Expressions (8) and (9) represent two different hypotheses relevant to the A main effect.

Jennings (1967) has suggested that a main effect can be defined in two different ways. Consider the following 2 X 2 design:

	B_1	B_2	
A_1	$\bar{X}_{11}$	$\bar{X}_{12}$	$\bar{X}_{1.}$
A_2	$\bar{X}_{21}$	$\bar{X}_{22}$	$\bar{X}_{2.}$

The symbol $\bar{X}_{11}$ represents the mean for group A_1B_1 while $\bar{X}_{1.}$ represents the overall mean for the A_1 row. The first method of defining the A main effect involves the difference between the overall row means ($\bar{X}_{1.}$ and $\bar{X}_{2.}$). The second definition of the A main effect involves the difference between the mean of the cell means in row 1, $(\bar{X}_{11} + \bar{X}_{12})/2$, and the mean of the cell means in row 2, $(\bar{X}_{21} + \bar{X}_{22})/2$. We will refer to these two definitions of main effect as the overall-mean main effect and the mean-of-means main effect. These two types of main effect are equivalent when cell frequencies are equal but differ from each other when cell frequencies are unequal.

Consider a 2 X 2 design with unequal n's. In order for the interaction to be residual to the overall-mean main effect for variable A, the following condition must hold:

$$\frac{1}{n_{i.}} \sum^{j} n_{ij} \alpha\beta_{ij} = 0, \text{ for all } i. \qquad (10)$$

If (10) is true, then the interaction components (the $\alpha\beta_{ij}$) do not influence the means ($\bar{X}_{1.}$ and $\bar{X}_{2.}$) defining the overall-mean A main effect. The hypothesis in (8) is a test of the overall-mean main effect. In order for the interaction to be residual to the mean-of-means main effect, $\overline{\alpha\beta}_{i.}$ must equal 0 for all i. Given $\overline{\alpha\beta}_{i.} = 0$, then the interaction effects do not influence the means, $(\bar{X}_{11} + \bar{X}_{12})/2$ and $(\bar{X}_{21} + \bar{X}_{22})/2$, defining the mean-of-means A main effect. The hypothesis in (9) is a test of the mean-of-means main effect.

Some writers--e.g., Winer (1971)--have suggested that analysis of main effects is identical for the equal-n case and the proportional-n case. The

present treatment suggests otherwise. In a proportional-n design, the main effect for a variable may be defined in two different ways--as an overall-mean main effect or a mean-of-means main effect. Which main effect a researcher tests depends upon the type of generalization he wishes to make. If his intention is to generalize to a population characterized by equal frequency of occurrence of the levels of B, then the mean-of-means A main effect should be tested. Alternatively, if his intention is to generalize to a population which is similar to the sample with regard to distribution of individuals across categories, then the overall-mean A main effect should be tested. Given this perspective, it is obvious why the two main effects are equivalent for equal-n designs but differ from each other for unequal-n designs.

Unbalanced (disproportional-n) designs with no empty cells. In disproportional-n designs, confounding exists with regard to all analysis of variance sources. In such designs, SS_A, $SS_{A/B}$, $SS_{A/AB}$, and $SS_{A/B,AB}$ can lead to four different values. Thus, F(A), F(A/B), F(A/AB), and F(A/B,AB) test different hypotheses. F(A) tests the following hypothesis:

$$\text{H: equality of } \alpha_i + \frac{1}{n_{i.}}\sum^{j} n_{ij}(\beta_j + \alpha\beta_{ij}) \quad \text{for all } i. \tag{11}$$

F(A) provides a test for the overall-mean A main effect confounded by the B effect.

The hypothesis in (11) may appear to be of little interest since it involves confounded sources. However, consider a researcher doing survey research on two categorical variables, A and B. Assume that A and B are confounded in the population being sampled and that B is much more expensive to measure than A. Because of the confounding in the population, unbalanced data will likely be obtained in the sample. The researcher may be mainly interested in the predictive power of the inexpensively measured A variable.

In this case, he may have included the B variable in the design simply to increase the power of the test of the A variable (by removing the B variation from the error term). It is his hope that A will prove a good predictor and that it will be unnecessary to measure B in the future. Given these circumstances, the hypothesis in (11) would be the hypothesis of interest. Such possible interest in (11) illustrates an important point. When data are unbalanced, interesting questions may not be limited to those questions involving unconfounded analysis of variance sources.

From page 310 of Searle (1971), it is observed that F(A/B) tests the following hypothesis in the two-factor case:

$$H: \quad (n_{i.} - \sum_{j} n_{ij}^2/n_{.j})\alpha_i - \sum_{i' \neq i} \sum_{j} (n_{ij}n_{i'j}/n_{.j})\alpha_{i'} + \sum_{j} (n_{ij} - n_{ij}^2/n_{.j})\alpha\beta_{ij} - \sum_{i' \neq i} \sum_{j} (n_{ij}n_{i'j}/n_{.j})\alpha\beta_{i'j} \quad (12)$$
$$= 0 \text{ for all } i = 1,2,\ldots,a-1 \text{ and for } i' \neq i.$$

It can be shown that when there are no empty cells and when $\sum_{i} n_{ij}\alpha\beta_{ij} = \sum_{j} n_{ij}\alpha\beta_{ij} = 0$, then the hypothesis in (12) reduces to the equality of the α_i for all i. (Appendix A presents a demonstration of this fact for the 2 X 2 case.) The hypothesis tested by F(A/B) may be thought of as a test of the unconfounded overall-mean main effect for variable A.

The authors have been unable to determine the hypothesis corresponding to F(A/AB). The hypothesis corresponding to F(A/B,AB), however, is known. This hypothesis is presented in (13).

$$H: \text{ equality of } \alpha_i + \overline{\alpha\beta}_{i.} \text{ for all } i. \quad (13)$$

Thus, F(A/B,AB) provides a test of the unconfounded mean-of-means main effect for variable A.

Unbalanced (disproportional-n) designs with empty cells. In the case of

designs involving empty cells, general expressions cannot be written for the hypotheses tested. What hypothesis is tested depends upon the pattern of the empty cells. Consider a design involving the following cell frequencies:

	B_1	B_2	B_3
A_1	1	0	2
A_2	0	1	1

While this design would never actually be analyzed because of the small number of observations, it will serve to demonstrate the type of hypothesis tested when empty cells are present.

With regard to the design above, F(A) tests the following hypothesis:

$$H: \alpha_1 - \alpha_2 + (\beta_1 + 2\beta_3)/3 - (\beta_2 + \beta_3)/2 + (\alpha\beta_{11} + 2\alpha\beta_{13})/3 - (\alpha\beta_{22} + \alpha\beta_{23})/2. \tag{14}$$

The dependence of (14) upon which cells are empty is readily apparent. If the cell frequencies in the design do not reflect the state of affairs in the population to which we wish to generalize, then (14) is of little interest.

For the design in question, F(A/B) tests the following hypothesis:

$$H: \alpha_1 - \alpha_2 + \alpha\beta_{13} - \alpha\beta_{23} = 0. \tag{15}$$

The hypothesis in (15) is particularly uninteresting. It does not provide an unconfounded test of the A main effect, whether that main effect is conceived of in terms of comparison of overall-means or comparison of means-of-means.

The F-ratio, F(A/B,AB), tests the following hypothesis with regard to the design in question:

$$H: \alpha_1 - \alpha_2 + (\alpha\beta_{11} + \alpha\beta_{13})/2 - (\alpha\beta_{22} + \alpha\beta_{23})/2. \tag{16}$$

This hypothesis represents a test of the mean-of-means main effect for the A variable if, and only if, the empty cells in the design reflect an absence of

the corresponding conditions (A_1B_2 and A_2B_1) in the population of interest.

Finally, the hypothesis tested by F(A/AB) for the design in question is not known to the authors. This concludes the discussion of the analysis of variance approach. Before proceeding to other approaches, a brief summary is in order.

Summary. The analysis of variance approach is different in application to different designs. For equal-n designs, analysis of variance offers a simple and conceptually powerful approach. The main effect hypotheses tested are clearly interesting and unambiguous. As designs deviate more and more from the equal-n variety, the application of analysis of variance is increasingly strained. For proportional-n designs, analysis of variance offers hypotheses concerning two different definitions of a main effect, and the researcher must decide which type of main effect is of interest. For disproportional-n designs, there are four statistical tests relevant to a given main effect. The hypotheses corresponding to these tests are often quite complex and difficult to interpret, especially in cases where empty cells are present.

Jennings' Regression Analysis Approach

The analysis of variance approach can be conceptualized as a theory oriented approach to data analysis. Observable differences (differences between means) are analyzed in an attempt to make inferences about hypothetical effects. When data are unbalanced, such inferences become difficult to make. In contrast, Jennings' (1967) approach to data analysis is more empirical in its orientation. The tested hypotheses relate directly to differences between means and no emphasis is placed upon making inferences about hypothetical effects. No particular problem arises in applying Jennings' method to unbalanced data, since the hypotheses of interest are easily tested on the basis of observable differences.

Jennings' approach involves definition of dummy variables corresponding to group membership. Group membership coding of dummy variables is illustrated in Table 4.

Table 4

Group Membership Coding for a 2X2 Design

Group	Z(1,1)	Z(1,2)	Z(2,1)	Z(2,2)
A_1B_1	1	0	0	0
A_1B_2	0	1	0	0
A_2B_1	0	0	1	0
A_2B_2	0	0	0	1

Note that one dummy variable is defined for each group in the design. Z(1,1) is the dummy variable corresponding to group A_1B_1 and Z(1,1) = 1 for subjects in group A_1B_1 but Z(1,1) = 0 for subjects not in group A_1B_1. The other dummy variables are defined in corresponding fashion.

Jennings' method of analysis involves four basic steps: (1) construction of a full model based upon group membership dummy variables, (2) statement of the hypothesis to be tested, (3) construction of a restricted model on the basis of this hypothesis, and (4) test of the hypothesis by comparing the full and restricted models.

Constructing the full model. The full model is a multiple regression equation relating the criterion variable to the dummy variables. For a 2x2 design, the full model is:

$$Y_{ijk} = a_1Z(1,1) + a_2Z(1,2) + a_3Z(2,1) + a_4Z(2,2) + e_{ijk}. \qquad (17)$$

In expression (17), Y_{ijk} is the criterion score for the k^{th} individual in the ij^{th} group; a_1, a_2, a_3, and a_4 are the regression coefficients for the dummy variables; and e_{ijk} is the error in prediction for the k^{th} individual in the ij^{th} group. The least squares estimates of the parameters in (17) are the arithmetic means for the groups. Thus,

$$\hat{a}_1 = \bar{Y}_{11.}$$
$$\hat{a}_2 = \bar{Y}_{12.}$$
$$\hat{a}_3 = \bar{Y}_{21.}$$
$$\hat{a}_4 = \bar{Y}_{22.}$$

where $\bar{Y}_{11.}$ is the mean for group A_1B_1; $\bar{Y}_{12.}$ is the mean for group A_1B_2; etc.

Statement of hypothesis. Hypotheses are couched in terms of the expected values for the different groups. Suppose one is interested in the question: Do subjects experiencing A_1 differ from subjects experiencing A_2? A null hypothesis corresponding to this question is

$$H: \bar{Y}_{11.} + \bar{Y}_{12.} = \bar{Y}_{21.} + \bar{Y}_{22.}. \tag{18}$$

The truth of this null hypothesis would imply the following restriction on the parameters of the full model:

$$a_1 + a_2 = a_3 + a_4$$

or solving for a_1,

$$a_1 = a_3 + a_4 - a_2. \tag{19}$$

Constructing the restricted model. The restricted model is constructed by incorporating the restriction (19) into the full model (17). Thus the restricted model for the hypothesis in expression (18) would be

$$Y_{ijk} = (a_3 + a_4 - a_2)Z(1,1) + a_2Z(1,2) + a_3Z(2,1) + a_4Z(2,2) + f_{ijk} \tag{20}$$

where f_{ijk} is the symbol for error.

Simplifying (20), we obtain

$$Y_{ijk} = a_2[Z(1,2) - Z(1,1)] + a_3[Z(1,1) + Z(2,1)] + a_4[Z(1,1) + Z(2,2)] + f_{ijk}. \tag{21}$$

From (21) it can be seen that there are three predictor variables in the restricted model and these predictor variables are linear transformations of the dummy variables in the full model.

Test of hypothesis. The hypothesis in (18) can be tested by comparing the full and restricted models. Let R(F) be the multiple correlation

coefficient calculated on the basis of the full model and R(R) be the multiple correlation coefficient calculated on the basis of the restricted model. The following F-ratio tests the hypothesis in question:

$$F = \frac{[R^2(F) - R^2(R)]/df_1}{[1 - R^2(F)]/df_2},$$

where df_1 = the number of unknown parameters in the full model minus the number of unknown parameters in the restricted model and df_2 = the number of subjects minus the number of unknown parameters in the full model.

Testable Hypotheses

In the 2x2 case, any hypothesis of the following form may be tested:

$$C_1\bar{Y}_{11.} + C_2\bar{Y}_{12.} = C_3\bar{Y}_{21.} + C_4\bar{Y}_{22.}, \qquad (22)$$

where the C_1, C_2, C_3, and C_4 coefficients are any constants. Consider the following two hypotheses:

$$H: \bar{Y}_{11.} + \bar{Y}_{12.} = \bar{Y}_{21.} + \bar{Y}_{22.} \qquad (23a)$$

$$\text{and } H: \frac{n_{11}}{n_{1.}}\bar{Y}_{11.} + \frac{n_{12}}{n_{1.}}\bar{Y}_{12.} = \frac{n_{21}}{n_{2.}}\bar{Y}_{21.} + \frac{n_{22}}{n_{2.}}\bar{Y}_{22.}. \qquad (23b)$$

Both of these hypotheses are testable hypotheses of the form (22). Hypothesis (23a) is analogous to the mean-of-means main effect (an equal-n population is of interest) discussed in connection with analysis of variance, while hypothesis (23b) corresponds to the overall-means main effect (a population with a distribution of individuals similar to the sample is of interest).

Another type of testable hypothesis should be considered. Assume a 2x2 design. Further assume that the population of interest is characterized by twice as many individuals in the A_1B_1 and A_2B_1 categories as in the A_1B_2 or A_2B_2 categories. Given intent to generalize to this population, the following hypothesis may be especially interesting:

$$2\bar{Y}_{11.} + \bar{Y}_{12.} = 2\bar{Y}_{21.} + \bar{Y}_{22.}. \qquad (24)$$

The coefficients in a hypothesis are relevant to population characteristics and not sample characteristics. The hypothesis in (24) can be meaningfully tested even if the cell frequencies in the sample do not reflect the distribution of individuals in the population. For example, this hypothesis can be tested with regard to an equal-n sample. Even if the distribution in the sample does not reflect the distribution in the population, Jennings' approach allows this problem to be easily skirted and allows tests to be made of hypotheses appropriate to the population. Such flexibility with regard to population of interest is an extremely important characteristic of Jennings' method.

Jennings' approach is easily applied to either balanced or unbalanced data. It is also applicable to designs involving empty cells. With this approach, any hypothesis based upon differences among means, calculable in the sample, can be tested and such an hypothesis can be couched in terms of different types of populations. A good label for Jennings' approach is the method of differences among means.

Other Approaches to Unbalanced Data Analysis

Several other techniques for analyzing unbalanced data have been suggested. Among these techniques are: (1) artificially balancing a design by discarding data or estimating missing data (Williams, 1972), (2) the unweighted means solution (Searle, 1971), (3) the method of expected subclass numbers (Steinhorst and Miller, 1969), and (4) the method of weighted squares of means (Searle, 1971). None of these techniques is very satisfying. The first three methods all involve approximate solutions which do not yield exact tests of the hypotheses of interest. While the method of weighted squares of means is not an approximate method (Searle, 1971), it tests no hypothesis which

cannot be tested by the previously discussed linear model techniques (analysis of variance and Jennings' regression analysis approach).

All of these other methods were first suggested in the 1930's and 1940's before advances in linear model theory had led to the development of exact procedures. One possible advantage for these other methods is that they are computationally simpler than the exact methods. However, with the present availability of computer facilities and statistical program packages, this advantage is obviated.

Discussion

Having dismissed the other alternatives, discussion can be limited to the analysis of variance approach and the differences among means approach (Jennings' approach). With balanced data, the analysis of variance approach provides clear tests of the statistical significance of the different analysis of variance sources (main effects and interactions). With unbalanced data, application of analysis of variance methods becomes less satisfying. When data are unbalanced, there are different manners for defining main effects. For disproportional-n data, inferences about unconfounded analysis of variance effects become difficult to make. In disproportional-n designs, simple observable differences (differences between means) confound the different analysis of variance effects. More complex observable differences (differences between means adjusted for the influence of covariables) allow unconfounded tests of significance for the different analysis of variance effects, as long as there are no empty cells. When empty cells are present, it may be impossible to test hypotheses concerning unconfounded analysis of variance effects.

Unbalanced data are often obtained in survey research because the variables of interest are confounded in nature. If a researcher's orientation

is practical, he may have little interest in teasing out the influence of unconfounded hypothetical effects. Rather, he may be more interested in applications based upon the study of the confounded effects which actually occur in nature. Data sets can delimit the types of questions which are appropriate to ask. The major focus in handling unbalanced data problems should not be determining how to ask the traditional analysis of variance questions about these data, but rather what interesting questions can be unambiguously asked and how these questions should be asked.

The differences among means (Jennings') approach to data analysis offers a valid alternative to the traditional analysis of variance approach--especially with regard to the analysis of unbalanced data. A basic characteristic of the differences among means approach is that easily answered questions are asked--questions which are easily understood and can be answered by reference to simple observable differences in the data. A major advantage of this approach is that the questions asked may be made relevant to different populations of interest.

While the analysis of variance approach and the differences among means approach offer different perspectives for the analysis of categorical data, both rely on tests of differences between means. Given this common ground, it should be possible to establish equivalences between hypotheses tested by the two approaches. Table 5 presents equivalences for a 2x2 design with disproportional-n's but no empty cells.

Table 5

Equivalent Hypotheses for a 2x2 Design with

Disproportional-n's and No Empty Cells

Analysis of Variance Test	Analysis of Variance Hypothesis	Differences Among Means Hypothesis (Jennings' Approach)
F(A)	equality of $\alpha_i + \sum_j n_{ij}(\beta_j + \alpha\beta_{ij})$ for all i	$\frac{n_{11}}{n_{1.}}\bar{Y}_{11.} + \frac{n_{12}}{n_{1.}}\bar{Y}_{12.} = \frac{n_{21}}{n_{2.}}\bar{Y}_{21.} + \frac{n_{22}}{n_{2.}}\bar{Y}_{22.}$
F(A/B)	equality of $\alpha_i + f(n_{ij}\alpha\beta_{ij})$ for all i	unknown
F(A/AB)	unknown	unknown
F(A/B,AB)	equality of $\alpha_i + \overline{\alpha\beta}_{i.}$ for all i	$\bar{Y}_{11.} + \bar{Y}_{12.} = \bar{Y}_{21.} + \bar{Y}_{22.}$

A few comments about Table 5 will conclude this paper. The unknown hypotheses in Table 5 are by no means unknowable; rather they are simply unknown to the present authors. The analysis of variance hypothesis for F(A/B) is the complex expression given as (12). For use in Table 5, we have simplified this expression to equality of $\alpha_i + f(n_{ij}\alpha\beta_{ij})$ for all i. The term, $f(n_{ij}\alpha\beta_{ij})$ refers to an unspecified function which has the property that if $\sum_i n_{ij}\alpha_{ij} = \sum_j n_{ij}\alpha\beta_{ij} = 0$ then $f = 0$.

References

Cohen, J. Multiple regression as a general data-analytic system. Psychological Bulletin, 1968, 70, 426-443.

Edwards, A. L. Experimental design in psychological research. New York: Holt, Rinehart & Winston, 1968.

Jennings, E. Fixed effects analyses of variance by regression analysis. Multivariate Behavioral Research, 1967, 2, 95-108.

Kerlinger, F. N., & Pedhazur, E. J. Multiple regression in behavioral research. New York: Holt, Rinehart & Winston, 1973.

Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks/Cole, 1968.

Marks, E. Analysis of variance techniques for unbalanced data. Review of Educational Research, 1974, 44, 351-364.

Overall, J. E., & Klett, C. J. Applied multivariate analysis. New York: McGraw-Hill, 1972.

Overall, J. E., & Spiegel, D. K. Concerning least squares analyses of experimental data. Psychological Bulletin, 1969, 72, 311-322.

Rao, C. R. Linear statistical inference and its applications. New York: Wiley, 1973.

Searle, S. R. Linear models. New York: Wiley, 1971.

Steinhorst, R. K., & Miller, C. D. Disproportionality of cell frequencies in psychological and educational experiments involving multiple classification. Educational and Psychological Measurement, 1969, 29, 799-811.

Williams, J. D. Two way fixed effects analysis of variance with disproportionate cell frequencies. Multivariate Behavioral Research, 1972, 1, 67-83.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

Appendix A

This Appendix provides A demonstration that, for A 2x2 design with no empty cells, the hypothesis tested by F(A/B) reduces to equality of α_i for All i if it is Assumed that $\sum_j n_{ij}\alpha\beta_{ij} = \sum_i n_{ij}\alpha\beta_{ij} = 0$.

The general form of the hypothesis tested by F(A/B) is given in the text As expression (12). For the specific case of A 2x2 with no empty cells, this hypothesis becomes:

$$H: \left(n_{1.} - \sum_j n_{1j}^2/n_{.j}\right)\alpha_1 - \sum_j (n_{1j}n_{2j}/n_{.j})\alpha_2 \qquad (a)$$

$$+\sum_j\left[(n_{1j} - n_{1j}^2/n_{.j})\alpha\beta_{1j}\right]$$

$$-\sum_j\left[(n_{1j}n_{2j}/n_{.j})\alpha\beta_{2j}\right] = 0.$$

Consider the last two terms in expression (a). Performing the summations in these terms, we have

$$\left[n_{11} - \frac{n_{11}^2}{n_{.1}}\right]\alpha\beta_{11} + \left[n_{12} - \frac{n_{12}^2}{n_{.2}}\right]\alpha\beta_{12}$$

$$-\left[\frac{n_{11}n_{21}}{n_{.1}}\right]\alpha\beta_{21} - \left[\frac{n_{12}n_{22}}{n_{.2}}\right]\alpha\beta_{22} \quad .$$

Rearranging terms:

$$n_{11}\alpha\beta_{11} + n_{12}\alpha\beta_{12} - \frac{n_{11}}{n_{\cdot 1}}\left[n_{11}\alpha\beta_{11} + n_{21}\alpha\beta_{21}\right]$$

$$- \frac{n_{12}}{n_{\cdot 2}}\left[n_{12}\alpha\beta_{12} + n_{22}\alpha\beta_{22}\right] \quad .$$

Rewriting in terms of summation notation:

$$\sum_j n_{1j}\alpha\beta_{1j} - \frac{n_{11}}{n_{\cdot 1}}\sum_i n_{i1}\alpha\beta_{i1} - \frac{n_{12}}{n_{\cdot 2}}\sum_i n_{i2}\alpha\beta_{i2} \quad . \qquad (b)$$

The expression in (b) clearly equals 0 if $\sum_i n_{ij}\alpha\beta_{ij} = \sum_j n_{ij}\alpha\beta_{ij} = 0$, and given such assumption, (a) reduces to

$$H: \left(n_{1\cdot} - \sum_j n_{1j}^2 / n_{\cdot j}\right)\alpha_1 - \sum_j (n_{1j}n_{2j}/n_{\cdot j})\alpha_2 \qquad (c)$$

$$= 0 \quad .$$

Performing the summations in (c):

$$H: \left[n_{1\cdot} - \left(\frac{n_{11}^2}{n_{\cdot 1}} + \frac{n_{12}^2}{n_{\cdot 2}}\right)\right]\alpha_1$$

$$- \left[\frac{n_{11}n_{21}}{n_{\cdot 1}} + \frac{n_{12}n_{22}}{n_{\cdot 2}}\right]\alpha_2 = 0 \quad .$$

Finding common denominators within the brackets:

$$\left[\frac{n_{1\cdot}n_{\cdot 1}n_{\cdot 2} - n_{11}^2 n_{\cdot 2} - n_{12}^2 n_{\cdot 1}}{n_{\cdot 1}\, n_{\cdot 2}}\right]\alpha_1 \qquad (d)$$

$$- \left[\frac{n_{11}n_{21}n_{\cdot 2} + n_{12}n_{22}n_{\cdot 1}}{n_{\cdot 1}\, n_{\cdot 2}}\right]\alpha_2 = 0 \quad .$$

Let c equal the expression for the coefficient of α_2. That is,

$$c = \frac{n_{11} n_{21} n_{\cdot 2} + n_{12} n_{22} n_{\cdot 1}}{n_{\cdot 1} n_{\cdot 2}} \quad . \qquad (e)$$

Rewriting (d):

$$\left[\frac{(n_{11}+n_{12}) n_{\cdot 1} n_{\cdot 2} - n_{11}^2 n_{\cdot 2} - n_{12}^2 n_{\cdot 1}}{n_{\cdot 1} n_{\cdot 2}} \right] \alpha_1 - c\alpha_2 = 0.$$

Rearranging terms:

$$\left[\frac{n_{11} n_{\cdot 1} n_{\cdot 2} - n_{11}^2 n_{\cdot 2} + n_{12} n_{\cdot 1} n_{\cdot 2} - n_{12}^2 n_{\cdot 1}}{n_{\cdot 1} n_{\cdot 2}} \right] \alpha_1 - c\alpha_2 = 0 ,$$

$$\left[\frac{n_{11} n_{\cdot 2} (n_{\cdot 1} - n_{11}) + n_{12} n_{\cdot 1} (n_{\cdot 2} - n_{12})}{n_{\cdot 1} n_{\cdot 2}} \right] \alpha_1 - c\alpha_2 = 0 ,$$

$$\left[\frac{n_{11} n_{\cdot 2} (n_{11} + n_{21} - n_{11}) + n_{12} n_{\cdot 1} (n_{12} + n_{22} - n_{12})}{n_{\cdot 1} n_{\cdot 2}} \right] \alpha_1 - c\alpha_2 = 0 .$$

Finally,

$$\left[\frac{n_{11} n_{21} n_{\cdot 2} + n_{12} n_{22} n_{\cdot 1}}{n_{\cdot 1} n_{\cdot 2}} \right] \alpha_1 - c\alpha_2 = 0.$$

The expression for the coefficient of α_1 is C -- see definition of c in expression (e). Therefore,

$$c\alpha_1 - c\alpha_2 = 0$$

And $\alpha_1 = \alpha_2$.

Thus, the hypothesis in (a) reduces to $\alpha_1 = \alpha_2$ when it is assumed that $\sum_i n_{ij}\alpha\beta_{ij} = \sum_j n_{ij}\alpha\beta_{ij} = 0$.